

Causal/Predictive Imperfect Channel State Information in Block-Fading Channels

Khoa D. Nguyen, *Member, IEEE*, Nick Letzepis, *Member, IEEE*,
Albert Guillén i Fàbregas, *Senior Member, IEEE*, Lars K. Rasmussen,
Senior Member, IEEE

Abstract

We consider a multi-input multi-output (MIMO) block-fading channel with a general model for channel state information at the transmitter (CSIT). The model covers systems with causal CSIT, where only CSIT of past fading blocks is available, and predictive CSIT, where CSIT of some future fading blocks is available. The optimal diversity-multiplexing tradeoff (DMT) and rate-diversity tradeoff (RDT) of the channel are studied under long-term power constraints. The impact of imperfect (mismatched) CSIT on the optimal DMT and RDT is also investigated. Our results show the outage diversity gain obtained by providing imperfect causal/predictive CSIT, leading to new insights into system design and analysis.

I. INTRODUCTION

The mitigation of fading is a particularly challenging aspect in the design of reliable and efficient wireless communication systems [1]. The methods available for dealing with fading-related impairments are influenced by many factors, where some of the most important are

K. D. Nguyen and N. Letzepis are with Institute for Telecommunications Research, University of South Australia, SPRI Building - Mawson Lakes Blvd., Mawson Lakes SA 5095, Australia, e-mail: {khoa.nguyen,nick.letzepis}@unisa.edu.au. Albert Guillén i Fàbregas is with the Department of Engineering, University of Cambridge, Trumpington street, Cambridge CB2 1PZ, UK, e-mail: guillen@ieee.org. Lars K. Rasmussen is with the Communication Theory Lab, School of Electrical Engineering and the ACCESS Linneaus Center, Royal Institute of Technology, Stockholm, Sweden, e-mail: lars.rasmussen@ieee.org.

This work was presented in part at International Symposium on Information Theory, Austin, Texas, June 2010.

This work has been supported by the Sir Ross and Sir Keith Smith Fund, Cisco Systems; the Australian Research Council under ARC grants RN0459498, DP0881160; the Royal Society International Travel Grant 2009/R2, 2009/R4; the Swedish Research Council under VR grant 621-2009-4666 and European Community's Seventh Framework Programme (FP7/2007-2013) /ERC grant agreement No 228044.

the time/frequency behaviour of the fading processes and system constraints in terms of delay and power. For systems with no delay constraints or systems subject to fast fading, the channel can be considered *ergodic*. In this case, long-interleaved fixed-rate codes that do not exceed the channel capacity can be employed to ensure an arbitrarily low probability of error [2, 3]. In contrast, for slow fading channels with delay constraints, a transmitted codeword may only experience a small finite number of independent fading realisations and hence the channel is *non-ergodic*.

The *block-fading* channel [2, 4] is a simple model that captures the essence of non-ergodic channels. Here, each codeword comprises a finite number of blocks, where each block experiences an independent fading realisation, which remains constant within a given block. In this case, the instantaneous input-output *mutual information* is a random variable dependent on the underlying fading distribution. For most fading statistics, the channel capacity is zero in the strict Shannon sense as there is a non-zero *outage probability* that a fixed information rate is not supported [2, 4]. The outage probability is the lowest achievable word error probability of codes with sufficiently long block length [5]. As such, a rate-reliability tradeoff exists, whereby for a fixed number of blocks, a high rate is penalised by a large error probability.

Most works that study the block-fading channel focus on adaptive transmission techniques in which the power and/or rate is adapted to the channel conditions subject to system constraints (see [6] for a recent review). Adaptation, however, requires a certain degree of knowledge of the channel fades, also referred to as *channel state information* (CSI), at the transmitter and receiver. While it is a common assumption that CSI is available at the receiver, the availability of CSIT is system dependent. Particularly, CSIT can be obtained through the reciprocal channel in time-division duplex systems [7], or via a dedicated feedback channel [8].

A large body of work considers full CSI at the transmitter (CSIT), i.e., the transmitter knows values of the fades on all blocks. Most notably, the works in [9–11] study systems with perfect CSIT, while systems with imperfect CSIT is analysed in [12, 13]. This approach has practical relevance for systems exhibiting a set of instantaneous parallel channels, such as Orthogonal Frequency Division Multiplexing (OFDM) systems. The full CSIT assumption provides an upper bound to the performance of delay-limited communications. However, there are practical situations where this assumption is invalid. Specifically, in time-varying fading channel, causality constraints impose that only CSIT of blocks up to the current block is available [14]. Further delay in acquiring the CSIT may impose strictly causal CSIT. In this

case, CSIT is only available with a delay of a few fading blocks. Strictly causal CSIT arises in systems that experience slow time-varying fading, where CSIT is obtained via a feedback, e.g. free-space optical systems [15]. Meanwhile, specific systems may allow the CSI of future fading blocks to be available at the transmitter. For example, in mobile communications over slowly spatial fading channels, the CSI of the channel at a future location of the mobile device can be obtained and made available at the transmitter¹.

In this paper, we analyse the outage performance of the multi-input multi-output (MIMO) block-fading channel with a general CSIT model, which includes systems with causal CSIT and systems where CSIT of future blocks is available. We consider both systems with perfect CSIT and mismatched CSIT. For systems with mismatched CSIT, the transmitter is provided with a noisy version of channel fading gains, modelled with a Gaussian distribution as in [13]. With perfect CSIT, power adaptation algorithms based on *dynamic programming* are proposed in [14, 16] for systems with causal CSIT. Generalisations of algorithms can be derived for systems with perfect strictly causal CSIT, as well as systems with future-block CSIT. However, dynamic programming does not provide much insight into the outage performance and may exhibit prohibitive complexity in many scenarios. With imperfect CSIT a feasible adaptive power allocation rule is not known, even in the full CSIT case [13]. We therefore study the asymptotic outage performance of the block-fading channel for various CSIT scenarios without explicitly solving the optimal power allocation problem. In particular, we derive the optimal diversity-multiplexing tradeoff (DMT) and the optimal rate-diversity tradeoff (RDT) of the block-fading channel with long-term average power constraints. From the tradeoffs we gain insights into the impact of causal and predictive CSIT, as well as of imperfect CSIT, on the asymptotic outage performance.

The analysis shows that reducing delays in obtaining the CSIT, or increasing the predictive CSIT to include additional future blocks, generally improves the DMT and RDT of the MIMO block-fading channel. Similarly, improving the quality of the CSIT generally provides large gains in outage diversity at any multiplexing gain/transmission rate. However, at a given multiplexing gain/transmission rate, the optimal DMT/RDT may be dominated by either the number of CSIT blocks available or the CSIT quality. Specifically, strictly causal CSIT provides gains in outage diversity only if the multiplexing gain/transmission rate is sufficiently

¹The CSIT of future fading blocks can also be predicted from the current channels realisation when the fading blocks are statistically dependent. However, correlated block-fading channels are not within the scope of this paper.

small. Furthermore, in agreement with the results in [17], the outage diversity of systems with strictly causal CSIT is always finite, even when the CSIT available is perfect. In contrast, systems with perfect predictive CSIT achieve infinite outage diversity, which in many cases leads to a positive delay-limited capacity [11, 18]. With imperfect predictive CSIT, increasing the number of predictive blocks improves the asymptotic outage performance, until the outage diversity is dominated by the CSIT noise. These results highlight the roles of CSIT and its quality on the asymptotic outage performance of the MIMO block-fading channel, thus providing guidelines for system design.

The remainder of the paper is organised as follows. The system model is described in Section II, while Section III provides some preliminaries necessary for the paper. In Section IV, the DMT of the block-fading channel with the optimal Gaussian input constellation is analysed, while Section V studies the RDT of MIMO block-fading channels with arbitrary discrete input constellations. Concluding remarks are given in Section VI and finally, proofs of various results are provided in the Appendices.

The following notations are used in the paper. Scalar variables are denoted with lowercase symbols, lowercase and uppercase boldface symbols correspondingly denote vector and matrix variables. Sets are denoted with calligraphic symbols, while real and complex sets are denoted with \mathbb{R} and \mathbb{C} correspondingly. Hermitian and non-conjugate transposes are correspondingly denoted by $(\cdot)^\dagger$ and $(\cdot)^T$. The trace of a matrix is denoted by $\text{tr}(\cdot)$; while $\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ denotes a block-diagonal matrix whose diagonal blocks are matrices $\mathbf{A}_1, \dots, \mathbf{A}_n$. Expectations are denoted by $\mathbb{E}[\cdot]$. The magnitude of ξ is denoted as $|\xi|$, while $\lfloor \xi \rfloor$ ($\lceil \xi \rceil$) denotes the largest (smallest) integer smaller (greater) than ξ .

II. SYSTEM MODEL

In this paper, we consider a MIMO block-fading channel, with N_τ transmit and N_v receive antennae. For convenience, define $n = \min\{N_\tau, N_v\}$ and $m = \max\{N_\tau, N_v\}$. Binary data is encoded with a code of rate R bits per channel use, constructed over an alphabet $\mathcal{X} \subseteq \mathbb{C}$. The resulting transmitted codeword consists of B blocks, where each block comprises of L vector channel uses of size $N_\tau \times 1$. We denote $\mathbf{x}_b[l] \in \mathcal{X}^{N_\tau \times 1}$ as the l th transmitted symbol vector of block b , for $l = 1, \dots, L$ and $b = 1, \dots, B$. The symbols are assumed to be drawn independently from \mathcal{X} with unit average energy, i.e., $\mathbb{E}[\mathbf{x}_b[l]\mathbf{x}_b^\dagger[l]] = \mathbf{I}_{N_\tau}$, where \mathbf{I}_N the $N \times N$ identity matrix.

We denote by \mathbf{H}_b the $N_v \times N_\tau$ complex channel matrix for block $b = 1, \dots, B$. These

matrices are drawn independently for each block, and remain fixed for the corresponding L channel uses. In addition, we assume the elements of the channel matrix are independently, identically distributed (i.i.d.) complex Gaussian random variables (the Rayleigh fading channel model [1])². Let $\widehat{\mathbf{H}}_b$, available at the transmitter, be a noisy version of the true channel realisation \mathbf{H}_b , so that

$$\mathbf{H}_b = \widehat{\mathbf{H}}_b + \mathbf{E}_b, \quad b = 1, \dots, B, \quad (1)$$

where $\mathbf{E}_b \in \mathbb{C}^{N_v \times N_\tau}$ is the CSIT noise matrix (independent of $\widehat{\mathbf{H}}_b$) whose entries are i.i.d. complex Gaussian random variables with zero mean and variance σ_e^2 . This model has been motivated for imperfect CSIT in many practical communication scenarios [13, 20–22]. As in [13, 22], we assume that the CSIT noise variance decays as

$$\sigma_e^2 = P^{-d_e} \quad (2)$$

for some $d_e \geq 0$, where P is the power constraint to be defined subsequently. For convenience, we introduce the normalised channel gains

$$\overline{\mathbf{H}}_b = \frac{\sqrt{2}}{\sigma_e} \mathbf{H}_b. \quad (3)$$

Given $\widehat{\mathbf{H}}_b$, then $\overline{\mathbf{H}}_b$ is a complex Gaussian matrix with mean $\frac{\sqrt{2}}{\sigma_e} \widehat{\mathbf{H}}_b$ and entries having a scaled unit variance (unit variance on real and imaginary dimensions).

At the transmission of block b , we assume the transmitter only has knowledge of $\widehat{\mathbf{H}}^{(b-u)} = \text{diag}(\widehat{\mathbf{H}}_1, \dots, \widehat{\mathbf{H}}_{b-u})$, where $-B \leq u \leq B$ is an arbitrary fixed integer. When $0 < u \leq B$, the parameter u models the delay in obtaining CSI at the transmitter, due to, e.g., propagation and processing delays. When $-B \leq u \leq 0$, the parameter u models the number of future blocks with predictive CSIT.

We assume that the signal at each receive antenna is corrupted by independent, zero-mean unit-variance additive white Gaussian noise (AWGN). Hence, under these assumptions, the b th block of $N_v \times L$ received noisy symbols is

$$\mathbf{Y}_b = \mathbf{H}_b \mathbf{P}_b^{\frac{1}{2}} \left(\widehat{\mathbf{H}}^{(b-u)} \right) \mathbf{X}_b + \mathbf{W}_b, \quad b = 1, \dots, B, \quad (4)$$

where $\mathbf{X}_b \in \mathcal{X}^{N_\tau \times L}$, $\mathbf{Y}_b \in \mathbb{C}^{N_v \times L}$; are correspondingly the transmit and receive signal in block b ; $\mathbf{W}_b \in \mathbb{C}^{N_v \times L}$ is the noise matrix whose elements are drawn i.i.d. from the zero-mean unit-variance Gaussian distribution; and $\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right) \in \mathbb{R}_+^{N_\tau \times N_\tau}$ is a diagonal matrix

²The Rayleigh fading assumption is included mainly for notational simplicity. The analysis can be extended to include a general fading distribution using the results in [19].

whose τ^{th} diagonal element denotes the power allocated to transmit antenna τ of block b . We further assume that the receiver has perfect knowledge of \mathbf{H}_b and $\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right)$ when receiving block b . The power allocation is subject to the long-term power constraint,

$$\mathbb{E} \left[\frac{1}{B} \sum_{b=1}^B \text{tr} \left(\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right) \right) \right] \leq P. \quad (5)$$

III. PRELIMINARIES

The channel described by (4) is not information stable under the assumption of quasi-static fading [23] and as a consequence, the capacity in the strict Shannon sense is zero. We therefore study the information outage probability,

$$\text{P}_{\text{out}}(P, R) = \Pr \left\{ \frac{1}{B} \sum_{b=1}^B I_{\mathcal{X}} \left(\mathbf{H}_b \mathbf{P}_b^{\frac{1}{2}} \left(\widehat{\mathbf{H}}^{(b-u)} \right) \right) < R \right\}, \quad (6)$$

which is a fundamental limit on the codeword error performance of any coding scheme [2, 4, 5]. In (6), $I_{\mathcal{X}}(\mathbf{S})$ denotes the input-output mutual information of a MIMO block-fading channel with input constellation \mathcal{X} and channel matrix \mathbf{S} . With the optimal Gaussian input constellation,

$$I_{\mathcal{X}}(\mathbf{S}) = \log_2 \det(\mathbf{I}_{N_v} + \mathbf{S} \mathbf{S}^\dagger); \quad (7)$$

while with a uniform discrete and fixed constellation \mathcal{X} of size 2^M ,

$$I_{\mathcal{X}}(\mathbf{S}) = MN_\tau - \frac{1}{2^{MN_\tau}} \sum_{\mathbf{x} \in \mathcal{X}^{N_\tau}} \mathbb{E}_{\mathbf{z}} \left[\log_2 \left(\sum_{\mathbf{x}' \in \mathcal{X}^{N_\tau}} e^{-\|\mathbf{S}(\mathbf{x} - \mathbf{x}') + \mathbf{z}\|^2 + \|\mathbf{z}\|^2} \right) \right]. \quad (8)$$

Given mismatched CSIT $\widehat{\mathbf{H}}^{(b-u)}$, $\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right)$ is the solution to the minimisation problem

$$\begin{cases} \text{minimise} & \text{P}_{\text{out}}(P, R) \\ \text{subject to :} & \mathbb{E} \left[\frac{1}{B} \sum_{b=1}^B \text{tr} \left(\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right) \right) \right] \leq P. \end{cases} \quad (9)$$

For systems with full perfect CSIT, i.e., $\mathbf{H}^{(B)}$ is known at the transmitter prior to transmission, the optimal power allocation rule and outage diversity is studied in [9–11]. With perfect causal CSIT, i.e., $\widehat{\mathbf{H}}^{(b-u)} = \mathbf{H}^{(b-u)}$, and $u = 0$, (9) can be solved via dynamic programming [14, 16]. The extension to $u > 0$ or $-B < u < 0$ is also possible, although the problem becomes exceedingly difficult as $|u|$ increases. With mismatched CSIT, the problem becomes even more challenging [13]. However, as we shall see, it is possible to examine the asymptotic behaviour of $\text{P}_{\text{out}}(P, R)$ without explicitly solving (9). In particular, for systems with a Gaussian input constellation, we study the DMT [24] $d(r)$ defined as

$$d(r) = \lim_{P \rightarrow \infty} \frac{-\log \text{P}_{\text{out}}(P, r \log_2 P)}{\log P}, \quad (10)$$

where $r \in [0, n]$ is the multiplexing gain. Meanwhile, for systems with discrete input constellation \mathcal{X} of size 2^M , we study the RDT $d(R)$,

$$d(R) \triangleq \lim_{P \rightarrow \infty} \frac{-\log P_{\text{out}}(P, R)}{\log P} \quad (11)$$

with $R \in (0, MN_\tau)$. Note that the optimal RDT has been derived in [13] for the special case $u = -B$, where mismatched CSIT of all fading blocks is known prior to the transmission of each codeword.

For systems with uniform power allocation and Gaussian input constellation, it follows from [24] that the DMT $d^{\text{uni}}(r)$ is the piecewise linear curve connecting the points $(k, B(N_\tau - k)(N_v - k))$, $k = 0, \dots, n$. Meanwhile, for systems with discrete input constellation \mathcal{X} of size 2^M , the RDT is given by the Singleton bound [6, 25]

$$d^{\text{uni}}(R) = N_v d_S(R) \triangleq N_v \left(1 + \left\lfloor B \left(N_\tau - \frac{R}{M} \right) \right\rfloor \right). \quad (12)$$

Note that $d^{\text{uni}}(R)$ is also the outage diversity of systems with short-term power constraint $\sum_{b=1}^B \text{tr} \left(\mathbf{P}_b \left(\widehat{\mathbf{H}}^{(b-u)} \right) \right) \leq BP$ [11]. The optimal DMT and RDT of systems with long-term power constraints are investigated in the subsequent sections.

IV. DIVERSITY-MULTIPLEXING TRADEOFF OF GAUSSIAN INPUT CHANNELS

A. Causal CSIT

When causal CSIT is available, the achievable outage diversity of a MIMO block-fading channel with long-term power constraint is given in the following theorem.

Theorem 1: Consider transmission over the MIMO block-fading channel in (4) with multiplexing gain $r \in [0, n]$. Assume that mismatched CSIT $\widehat{\mathbf{H}}^{(b-u)}$ as modelled in (1) is available at the transmission of block b , where $u > 0$ is the delay in obtaining CSIT. Then the optimal DMT is lower bounded by

$$d(r, d_e) \geq \min_{\mathbf{k} \in \{0, \dots, n\}^B} d_{\mathbf{k}}, \quad (13)$$

where $d_{\mathbf{k}}$ is given as

$$\left\{ \begin{array}{ll} \text{Infimum} & Bmnd_e + \sum_{b=1}^B \sum_{i=1}^n (2i - 1 + m - n) \bar{\omega}_{b,i} \\ \text{Subject to} & \sum_{b=1}^B \sum_{i=1}^n \left(\pi_b \left(\widehat{\omega}^{(b-u)} \right) - d_e - \bar{\omega}_{b,i} \right)_+ < Br \\ & \bar{\omega}_{b,1} \geq \dots \geq \bar{\omega}_{b,n-k_b} \geq 0 \geq \bar{\omega}_{b,n-k_b+1} \geq \dots \geq \bar{\omega}_{b,n} \geq -d_e \\ & \pi_b \left(\widehat{\omega}^{(b-u)} \right) = 1 + mnd_e(b-u)_+ + \sum_{b'=1}^{b-u} \sum_{i=n-k_{b'}+1}^n (2i - 1 + m - n) \bar{\omega}_{b',i} \end{array} \right. \quad (14)$$

Proof: See appendix A-A. ■

The optimisation problem in (14) is linear, and can therefore easily be solved by a simplex algorithm³. The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$ and causal CSIT with delay $u = 3$ is illustrated in Figure 1. The figure shows that the optimal DMT of a MIMO block-fading can be improved by increasing the quality of causal CSIT. However, causal CSIT does not provide any gain in outage diversity for multiplexing gains $r \geq 1.25$. Outage diversity gains at high multiplexing gains r are only observed in systems with small u , as illustrated in Figure 2. Figure 2 shows the achievable DMT of the same MIMO block-fading channel ($N_\tau = N_v = 2, B = 4$) with perfect causal CSIT for various delay u , illustrating the significant impact of CSIT delay on the asymptotic outage performance. Note in Figure 1 that for a given CSIT delay u , a finite d_e is sufficient to achieve the optimal outage diversity of a corresponding system with perfect CSIT. For example, with $u = 3$, $d_e = 0.5$ exhibits the same DMT as systems with $d_e = \infty$ for multiplexing gains $r \geq 0.5$; also note that $d_e = 1$ is sufficient in terms of outage diversity for all $r \in [0, n]$. The d_e threshold required for obtaining the optimal DMT of a perfect causal CSIT will be analytically shown in the sequel for the vector channel case.

Due to the complexity of the optimisation problem in (14), the impact of the various parameters such as d_e and u on the DMT curve is difficult to deduce in general. We therefore consider the vector channel case ($n = 1$) as given in the following.

Corollary 1: Consider transmission with multiplexing gain $r \in [0, 1]$ over the block-fading channel in (4) where $n = 1$. Assume that causal mismatched CSIT $\widehat{\mathbf{H}}^{(b-u)}$ as modelled in (1) is available at the transmission of block b , where u is the delay in obtaining CSIT. The optimal DMT is lower bounded by⁴

$$d(r, d_e) \geq \begin{cases} mB(1-r), & B - u - Br \leq 0 \\ m \sum_{i=1}^{B-\lfloor Br \rfloor} a_i^*, & \text{otherwise.} \end{cases} \quad (15)$$

³The problem is solved by first performing a minimisation, with inequalities constraints $<$ replaced by \leq , then the discontinuous points (if any) are taken care of based on the solution obtained.

⁴In this case, it can be shown that the following achievable DMT is optimal following the same steps as in Theorem 1. Here the outage diversity is simpler to characterize and the converse is obtained since the argument $\mathbf{S}\mathbf{S}^\dagger$ of the mutual information expression in (7) reduces to a scalar value.

where a_i^* 's are defined as follows,

$$a_i^* = \begin{cases} 1 - Br + \lfloor Br \rfloor, & i = 1 \\ 1, & i = 2, \dots, u \\ a_{i-1}^* + m \min \{a_{i-u}^*, d_e\} & i = u + 1, \dots, B - \lfloor Br \rfloor \end{cases} \quad (16)$$

Proof: See Appendix A-B. ■

The expression of $d(r, d_e)$ in (15) confirms the thresholds observed in Figure 1 for the vector channel case. In particular, no gain in outage diversity is obtained by causal CSIT if the delay u satisfies $u \geq B(1 - r)$. Furthermore, it also follows from (15) and (16) that increasing d_e beyond $a_{B-\lfloor Br \rfloor - u}^*$ does not increase the outage diversity gain. Equivalently, instead of having perfect CSIT, a system whose CSIT error decays with P as $P^{-a_{B-\lfloor Br \rfloor - u}^*}$ is sufficient in terms of outage diversity.

B. Predictive CSIT

For systems with predictive CSIT, where mismatched CSIT of blocks up to $b+t$ is available at the transmission of block b , the achievable DMT is given as follows.

Theorem 2: Consider transmission over the MIMO block-fading channel in (4) with multiplexing gain $r \in [0, n]$. Assume that predictive mismatched CSIT $\widehat{\mathbf{H}}^{(b+t)}$ as modelled in (1) is available at the transmission of block b , where $t \geq 0$ is the number of CSIT blocks predicted. Then the optimal outage diversity is lower bounded by

$$d(r, d_e) \geq \min_{\mathbf{k} \in \{0, \dots, n\}^B} d_{\mathbf{k}}, \quad (17)$$

where $d_{\mathbf{k}}$ is obtained by

$$\begin{cases} \text{Infimum} & \sum_{b=1}^B d_e(m - k_b)(n - k_b) + \sum_{b=1}^B \sum_{i=1}^{n-k_b} (2i - 1 + m - n) \bar{\omega}_{b,i} \\ \text{Subject to} & \sum_{b=1}^B \sum_{i=1}^{n-k_b} \left(\pi_b \left(\widehat{\omega}^{(b+t)} \right) - d_e - \bar{\omega}_{b,i} \right)_+ + \sum_{b=1}^B k_b \pi_b(\mathbf{k}) < Br \\ & \bar{\omega}_{b,i} \geq 0, i = 1, \dots, n - k_b, \end{cases} \quad (18)$$

where $\pi_b(\mathbf{k}) = 1 + d_e \sum_{b'=1}^{\min\{B, b+t\}} (m - k_{b'})(n - k_{b'})$.

Proof: See Appendix A-C. ■

The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$ is illustrated in Figures 3 and 4. Figure 3 clearly demonstrates the benefits of predicting CSIT. With $t = 0$, i.e., when only CSIT of the current transmission block is known, significant gains in outage diversity is observed, as already pointed out in [26]. Further gains in outage diversity are

possible by increasing t up to $B - 1$. Note that comparing to a system with t predictive blocks, predicting $t + 1$ fading blocks provides additional CSIT information only for the first $B - t - 1$ transmission blocks. Therefore, the additional outage diversity gains offered by predicting $t + 1$ blocks compared to that of predicting t blocks decreases with t , as observed in Figure 3. An alternative way to improve the outage diversity is to provide better CSIT, as illustrated in Figure 4. In agreement with the results in [26], increasing d_e significantly improves the outage diversity. In fact, even with $d_e = 1$, the outage diversity is so large that it can be considered infinite for all practical purposes. In contrast to systems with causal CSIT, $d(r, d_e)$ for systems with predictive CSIT is strictly increasing with d_e . With $d_e \rightarrow \infty$, as in systems where the CSIT quality grows exponentially with SNR, $d(r, d_e)$ reaches infinity, leading to a positive delay-limited capacity in many scenarios [18].

The theorem shows the impact predictive CSIT and its quality on the asymptotic outage performance, leading to essential system design guidelines. Particularly, for systems with limited resources for channel estimation, it may be better to have high quality CSIT for a few future blocks, rather than predicting far into the future with low quality estimations.

V. RATE-DIVERSITY TRADEOFF OF DISCRETE INPUT CHANNELS

In this section, we concentrate on the more practical case where the input constellation is discrete with a fixed and finite constellation \mathcal{X} of size 2^M . Systems with causal CSIT are studied in Section V-A, and then systems with predictive CSIT is studied in Section V-B.

A. Causal CSIT

Theorem 3 (Causal mismatched CSIT): Consider transmission at rate $R \in [0, N_\tau M]$ over the MIMO block-fading channel given in (4) with input constellation \mathcal{X} of size 2^M . Assume that mismatched CSIT $\widehat{\mathbf{H}}^{(b-u)}$ as modelled in (1) is available at transmission block b , where $u > 0$ is the delay in obtaining the CSIT. With the long-term power constraint in (5), the optimal RDT is given by

$$d(R, d_e) = N_v N_\tau \sum_{b=1}^{\hat{b}} a_b + N_v (d_S(R) - \hat{b} N_\tau) a_{\hat{b}+1}, \quad (19)$$

where

$$d_S(R) = 1 + \left\lfloor B \left(N_\tau - \frac{R}{M} \right) \right\rfloor \quad (20)$$

$$\hat{b} = \left\lfloor \frac{d_S(R)}{N_\tau} \right\rfloor \quad (21)$$

$$a_b = \begin{cases} 1, & b = 1, \dots, u \\ a_{b-1} + N_\tau N_v \min \{d_e, a_{b-u}\}, & b = u+1, \dots, \hat{b}+1. \end{cases} \quad (22)$$

Proof: See Appendix B-A. ■

The optimal rate-diversity tradeoff for a MIMO block-fading channel with $B = 4$, $N_\tau = N_v = 2$ is illustrated in Figures 5 and 6. Figure 5 shows that the optimal outage diversity increases significantly with decreasing delay in getting CSIT. Note that similarly to the Gaussian input case, even perfect causal CSIT may not provide gains in terms of outage diversity over the non-CSIT case when the transmission rate is sufficiently large. The observation can be explained from Theorem 3 as follows. The outage diversity $d(R, d_e)$ coincides with that of a system with no CSIT when the a_b 's in (19) are equal to 1. Therefore, causal CSIT provides gains in terms of outage diversity if and only if $d_S(R) > uN_\tau$, or equivalently when

$$R \leq \frac{B-u}{B} M N_\tau. \quad (23)$$

Meanwhile, Figure 6 shows that for a given delay u , significant gains in outage diversity can be obtained by improving the quality of CSIT (increasing d_e). Note from Theorem 3 that the outage diversity does not improve with increasing d_e when $d_e \geq a_{\lceil \frac{d_S(R)}{N_\tau} \rceil - u}$. In other words the optimal outage diversity of systems with perfect CSIT can be achieved with finite d_e . This agrees, and generalises, the result in [17], which shows that an ARQ system with a finite number of feedback bits can achieve the optimal outage diversity of that with infinitely many feedback bits. The result is numerically illustrated in Figure 6 for a 2-by-2 MIMO block-fading channel with $B = 4$, $u = 2$ and 16-QAM input constellation, where $d_e \geq 1$ is sufficient to achieve the outage diversity of perfect CSIT systems for all transmission rates.

B. Predictive CSIT

When predictive CSIT $\widehat{\mathbf{H}}^{(b+t)}$, for some $t \geq 0$, is available at transmission of block b , the optimal RDT is determined as follows.

Theorem 4 (Predictive mismatched CSIT): Consider transmission at rate R over the MIMO block-fading channel in (4) using input constellation \mathcal{X} of size 2^M . Assume that mismatched

CSIT $\widehat{\mathbf{H}}^{(b+t)}$ as modelled in (1) is available at the transmission of block b , where $t \geq 0$ is the number of future blocks with CSIT. With the long-term power constraint in (5), the optimal RDT is

$$d(R, d_e) = \begin{cases} N_v d_S(R)(1 + N_v d_S(R) d_e), & t \geq \hat{b} \\ N_v \left(d_S(R) + N_v d_e \left(\frac{(\hat{b}-t)(\hat{b}+t+1)}{2} N_\tau^2 + d_S(R) \left(d_S(R) - N_\tau(\hat{b}-t) \right) \right) \right), & \text{otherwise,} \end{cases} \quad (24)$$

where $d_S(R) = 1 + \lfloor B(N_\tau - \frac{R}{M}) \rfloor$ and $\hat{b} = \lfloor \frac{d_S(R)}{N_\tau} \rfloor$.

Proof: See Appendix B-B. ■

Theorem 4 illustrates the impact of mismatched predictive CSIT on the outage diversity of the MIMO block-fading channel. In contrast to the causal CSIT case, we observe from (24) that the optimal outage diversity is strictly increasing with the quality of CSIT d_e . In effect, $d(R, d_e) = \infty$ with perfect CSIT ($d_e \rightarrow \infty$). Moreover, for SISO systems with $t \geq 1$, or MIMO systems with $t \geq 0$, the outage curve is vertical [10, 11] when $d_e = \infty$, leading to positive delay-limited capacity [18]. The effect of d_e on the outage diversity of a 2-by-2 MIMO block-fading channel with $B = 4$ is illustrated in Figure 7. The figure shows that significant gains in outage diversity is obtained even for $t = 0$ and relatively small d_e , making the outage diversity effectively infinite for practical purposes, especially for small transmission rates.

Similarly, the outage diversity is improved by increasing t , the number of blocks whose CSIT is available prior to transmission. The rate-diversity tradeoff for a 2-by-2 MIMO block-fading channel with $B = 4$ using 16-QAM input constellation and $d_e = 0.5$ is illustrated in Figure 8. Note from (24) that the outage diversity cannot be further improved by increasing t beyond $\hat{b} = \lfloor \frac{d_S(R)}{N_\tau} \rfloor$. This effect is illustrated in Figure 8, where increasing t only improves the outage diversity at lower transmission rates. At $t = B - 1$, CSIT of all blocks is available prior to transmission, and thus the rate-diversity tradeoff curve coincides with that of systems with full mismatched CSIT.

Similarly to the Gaussian case in Section IV-B, with large t , increasing the number of predictive blocks leads to marginal improvement in outage diversity. Therefore, for systems with limited resources for channel estimation, it may be more beneficial to have high quality predictive CSIT for a few future blocks, rather than having poor predictions for many future blocks. Therefore, trading off between the number predictive blocks and the CSIT quality is required to effectively exploit the available channel-estimation resources.

VI. CONCLUSIONS

We have studied the asymptotic outage performance of the MIMO block-fading channel with a general model for incomplete CSIT. The model covers a wide range of scenarios, including systems where CSIT of all fading blocks is known prior to transmission, systems with causal CSIT where a delay in obtaining CSIT is incurred, and systems with predictive CSIT where the fading gains of future fading blocks is made available at the transmitter. The results illustrate the effects of the limited as well as imperfect CSIT on the optimal DMT and RDT under long-term average power constraints. The analysis reveals that the DMT (RDT) of systems with causal CSIT is limited by the delay in obtaining CSIT. Meanwhile, the DMT(RDT) of systems with predictive CSIT is limited by the quality of CSIT. Therefore, the quality and quantity (CSIT delay or number of predictive blocks) tradeoff is dependent on the type of CSIT available, leading to different design criteria in acquiring CSIT.

APPENDIX A

DMT OF MIMO BLOCK-FADING CHANNELS WITH MISMATCHED CSIT

A. Causal CSIT– Proof of Theorem 1

For a MIMO block-fading channel in (4), let $\lambda_b = [\lambda_{b,1}, \dots, \lambda_{b,n}]^T$, where $0 \leq \lambda_{b,1} \leq \dots \leq \lambda_{b,n}$ are the ordered eigenvalues of $\mathbf{H}_b \mathbf{H}_b^\dagger$; and let $\hat{\lambda}_b = [\hat{\lambda}_{b,1}, \dots, \hat{\lambda}_{b,n}]^T$, where $0 \leq \hat{\lambda}_{b,1} \leq \dots \leq \hat{\lambda}_{b,n}$ are the ordered eigenvalues of $\hat{\mathbf{H}}_b \hat{\mathbf{H}}_b^\dagger$. For an achievability result, we consider the power allocation rule $P_b(\hat{\mathbf{H}}^{(b-u)}) = P_b(\hat{\lambda}^{(b-u)}) \mathbf{I}_{N_\tau}$, where $\hat{\lambda}^{(b-u)} \triangleq (\hat{\lambda}_1, \dots, \hat{\lambda}_{b-u})$. Then, the outage probability asymptotically achieves

$$P_{\text{out}}(P, r \log_2 P) \doteq \Pr \left\{ \sum_{b=1}^B \sum_{i=1}^n \log_2 \left(1 + P_b(\hat{\lambda}^{(b-u)}) \lambda_{b,i} \right) < Br \log_2 P \right\}. \quad (25)$$

Let $\hat{\omega}_{b,i} = \frac{-\log \hat{\lambda}_{b,i}}{\log P}$, following [24], the distribution of $\hat{\omega}_b = [\hat{\omega}_{b,1}, \dots, \hat{\omega}_{b,n}]^T$ in the limit of large P is

$$f_{\hat{\omega}_b}(\hat{\omega}_b) = \begin{cases} \prod_{i=1}^n P^{-(2i-1+m-n)\hat{\omega}_{b,i}}, & \hat{\omega}_{b,1} \geq \dots \geq \hat{\omega}_{b,n} \geq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

Let $\hat{\omega}^{(b-u)} = [\hat{\omega}_1, \dots, \hat{\omega}_{b-u}]$ and $\pi_b(\hat{\omega}^{(b-u)}) = \frac{\log P_b(\hat{\lambda}^{(b-u)})}{\log P}$. The power allocation rule asymptotically satisfies

$$\int_{\hat{\omega}^{(b-u)} \in \mathbb{R}_+^{(b-u)n} : \hat{\omega}_{b',i} \geq \hat{\omega}_{b',i+1}} P^{\pi_b(\hat{\omega}^{(b-u)})} \prod_{b'=1}^{b-u} \prod_{i=1}^n P^{-(2i-1+m-n)\hat{\omega}_{b',i}} d\hat{\omega}^{(b-u)} \leq P. \quad (27)$$

Following the Varadhan's lemma [27, Sec. 4.3], the power constraint is asymptotically equivalent to

$$\pi_b \left(\hat{\omega}^{(b-u)} \right) \leq 1 + \sum_{b'=1}^{b-u} \sum_{i=1}^n (2i-1+m-n) \hat{\omega}_{b',i}. \quad (28)$$

Since the outage probability is a decreasing function of transmit power, the power allocation rule with

$$\pi_b \left(\hat{\lambda}^{(b-u)} \right) \equiv \pi_b \left(\hat{\omega}^{(b-u)} \right) = 1 + \sum_{b'=1}^{b-u} \sum_{i=1}^n (2i-1+m-n) \hat{\omega}_{b',i} \quad (29)$$

is optimal in terms of outage exponent.

Therefore, letting $\omega_{b,i} \triangleq \frac{-\log \lambda_{b,i}}{\log P}$ ($b = 1, \dots, B, i = 1, \dots, n$), it follows from (25) that the outage probability at large SNR behaves like

$$P_{\text{out}}(P, r \log_2 P) \doteq \Pr \left\{ \sum_{b=1}^B \sum_{i=1}^n \log_2 \left(1 + P^{\pi_b(\hat{\omega}^{(b-u)}) - \omega_{b,i}} \right) < Br \log_2 P \right\} \quad (30)$$

$$\doteq \Pr \left\{ \sum_{b=1}^B \sum_{i=1}^n \left(\pi_b \left(\hat{\omega}^{(b-u)} \right) - \omega_{b,i} \right)_+ < Br \right\}. \quad (31)$$

Now for $b = 1, \dots, B$, let $\bar{\lambda}_b = (\bar{\lambda}_{b,1}, \dots, \bar{\lambda}_{b,n})^T$, where $\bar{\lambda}_{b,1} \leq \dots \leq \bar{\lambda}_{b,n}$ are the ordered eigenvalues of $\bar{\mathbf{H}}_b \bar{\mathbf{H}}_b^\dagger$. Furthermore, letting $\bar{\omega}_{b,i} \triangleq \frac{-\log \bar{\lambda}_{b,i}}{\log P}$ (for $b = 1, \dots, B, i = 1, \dots, n$), it follows from (1) and (3) that $\bar{\omega}_{b,i} = -\omega_{b,i} + d_e$. Then the outage probability can be written as

$$P_{\text{out}}(P, r \log_2 P) \doteq \int_{(\hat{\omega}^{(B)}, \bar{\omega}^{(B)}) \in \mathcal{O}} \prod_{b=1}^B f_{\bar{\lambda}_b | \hat{\lambda}_b}(\bar{\lambda}_b | \hat{\lambda}_b) f_{\hat{\lambda}_b}(\hat{\lambda}_b) d\hat{\lambda}^{(B)} d\bar{\lambda}^{(B)}, \quad (32)$$

where

$$\mathcal{O} \triangleq \left\{ \left(\hat{\omega}^{(B)}, \bar{\omega}^{(B)} \right) \in (\mathbb{R}^{Bn}, \mathbb{R}^{Bn}) : \sum_{b=1}^B \sum_{i=1}^n \left(\pi_b \left(\hat{\omega}^{(b-u)} \right) - \bar{\omega}_{b,i} - d_e \right)_+ < Br \right\} \quad (33)$$

is the outage set.

Following the analysis in [26], the outage probability is bounded by

$$\begin{aligned} P_{\text{out}}(P, r \log_2 P) &\leq \int_{\mathcal{O}} \prod_{b=1}^B \prod_{i=1}^n \exp \left(-P^{-\bar{\omega}_{b,i}} - P^{-(\hat{\omega}_{b,i} - d_e)} + P^{\frac{-\bar{\omega}_{b,i} - (\hat{\omega}_{b,i} - d_e)}{2}} \right) P^{-(m-n)\bar{\omega}_{b,i}} \\ &\quad \prod_{j>i} \left[(P^{-\bar{\omega}_{b,i}} - P^{-\bar{\omega}_{b,j}})^2 \right] P^{-(m-n)\hat{\omega}_{b,i}} \prod_{j>i} \left[(P^{-\hat{\omega}_{b,i}} - P^{-\hat{\omega}_{b,j}})^2 \right] \\ &\quad \exp \left(-P^{-\hat{\omega}_{b,i}} \right) P^{-(\bar{\omega}_{b,i} + \hat{\omega}_{b,i})} d\hat{\omega}^{(B)} d\bar{\omega}^{(B)} \end{aligned} \quad (34)$$

As in [26], defining the $(n+1)^B$ disjoint integral regions

$$\begin{aligned}\mathcal{A}_{\mathbf{k}} &\triangleq \{\bar{\omega}^{(B)}, \hat{\omega}^{(B)} : \hat{\omega}_{b,1} \geq \dots \geq \hat{\omega}_{b,n-k_b} \geq d_e > \hat{\omega}_{b,n-k_b+1} \geq \dots \geq \hat{\omega}_{b,n} \geq 0, \\ &\quad \bar{\omega}_{b,1} \geq \dots \geq \bar{\omega}_{b,n-k_b} \geq 0, \\ &\quad \bar{\omega}_{b,n-k_b+1} = \hat{\omega}_{b,n-k_b+1} - d_e, \dots, \bar{\omega}_{b,n} = \hat{\omega}_{b,n} - d_e, b = 1, \dots, B\} \quad (35)\end{aligned}$$

where $\mathbf{k} = [k_1, \dots, k_B] \in \{0, \dots, n\}^B$. Further define the corresponding exponent $d_{\mathbf{k}}$ such that

$$\int_{\mathcal{O} \cap \mathcal{A}_{\mathbf{k}}} \prod_{b=1}^B f_{\bar{\omega}_b | \hat{\omega}_b}(\bar{\omega}_b | \hat{\omega}_b) f_{\hat{\omega}_b}(\hat{\omega}_b) d\bar{\omega}_b d\hat{\omega}_b \doteq P^{-d_{\mathbf{k}}}. \quad (36)$$

Then, the outage diversity at multiplexing gain r satisfies

$$d(r, d_e) \geq \min_{\mathbf{k}} \{d_{\mathbf{k}}\}. \quad (37)$$

We now have that [26]

$$\begin{aligned}\int_{\mathcal{O} \cap \mathcal{A}_{\mathbf{k}}} \prod_{b=1}^B f_{\bar{\omega}_b | \hat{\omega}_b}(\bar{\omega}_b | \hat{\omega}_b) f_{\hat{\omega}_b}(\hat{\omega}_b) d\bar{\omega}_b d\hat{\omega}_b &\doteq \\ \int_{\mathcal{O} \cap \mathcal{A}_{\mathbf{k}}} \prod_{b=1}^B P^{-\sum_{i=1}^{n-k_b} (2i-1+m-n)\bar{\omega}_{b,i}} P^{-\sum_{i=1}^n (2i-1+m-n)\hat{\omega}_{b,i}} d\bar{\omega}^{(B)} d\hat{\omega}^{(B)}. \quad (38)\end{aligned}$$

Together with (33), we have that

$$d_{\mathbf{k}} = \inf_{\mathcal{A}_{\mathbf{k}}} \left\{ \sum_{b=1}^B \sum_{i=1}^{n-k_b} (2i-1+m-n)\bar{\omega}_{b,i} + \sum_{i=1}^n (2i-1+m-n)\hat{\omega}_{b,i} \right\} \quad (39)$$

$$\text{s.t. } \sum_{b=1}^B \sum_{i=1}^{n-k_b} \left(\pi_b \left(\hat{\omega}^{(b-u)} \right) - d_e - \bar{\omega}_{b,i} \right)_+ + \sum_{i=n-k_b+1}^n \left(\pi_b \left(\hat{\omega}^{(b-u)} \right) - \hat{\omega}_{b,i} \right)_+ < Br, \quad (40)$$

where we recall that $\pi_b \left(\hat{\omega}^{(b-u)} \right) = 1 + \sum_{b'=1}^{b-u} \sum_{j=1}^n (2j-1+m-n)\hat{\omega}_{b',i}$. For $\hat{\omega}_{b,i}$ with $i \leq n-k_b$, decreasing $\hat{\omega}_{b,i}$ decreases the objective function, while the constraint is unchanged.

Therefore, the optimiser satisfies $\hat{\omega}_{b,i} = d_e, i = 1, \dots, n-k_b$. Thus, $d_{\mathbf{k}}$ is obtained from

$$\begin{cases} \text{Infimum} & Bmnd_e + \sum_{b=1}^B \sum_{i=1}^n (2i-1+m-n)\bar{\omega}_{b,i} \\ \text{Subject to} & \sum_{b=1}^B \sum_{i=1}^n \left(\pi_b \left(\hat{\omega}^{(b-u)} \right) - d_e - \bar{\omega}_{b,i} \right)_+ < Br \\ & \bar{\omega}_{b,1} \geq \dots \geq \bar{\omega}_{b,n-k_b} \geq 0 \geq \bar{\omega}_{b,n-k_b+1} \geq \dots \geq \bar{\omega}_{b,n} \geq -d_e \end{cases} \quad (41)$$

where $\pi_b \left(\hat{\omega}^{(b-u)} \right) = 1 + mnd_e(b-u)_+ + \sum_{b'=1}^{b-u} \sum_{i=n-k_b'+1}^n (2i-1+m-n)\bar{\omega}_{b',i}$. ■

B. Causal CSIT in Vector Channels– Proof of Corollary 1

With $n = 1$, letting $a_b = \bar{\omega}_{b,1} + d_e$. Noting that the solution of (14) satisfies $-d_e \leq \bar{\omega}_{b,1} \leq \pi_b(\hat{\omega}^{(b-u)}) - d_e$, a lower bound of $d(r, d_e)$ is obtained from Theorem 1 as

$$\begin{cases} \text{Infimum} & m \sum_{b=1}^B a_b \\ \text{Subject to} & \sum_{b=1}^B (\pi_b(\mathbf{a}) - a_b) < Br \\ & \pi_b(\mathbf{a}) \geq a_b \geq 0, b = 1, \dots, B, \end{cases} \quad (42)$$

where $\pi_b(\mathbf{a}) = 1 + m \sum_{b'=1}^{b-u} \min\{a_{b'}, d_e\}$. The constraints in (42) are equivalent to

$$\begin{cases} \sum_{b=1}^B a_b > \sum_{b=1}^B \pi_b(\mathbf{a}) - Br. \\ \pi_b(\mathbf{a}) \geq a_b \geq 0, b = 1, \dots, B. \end{cases} \quad (43)$$

We consider the following cases.

- When $B - u - Br \leq 0$, since $\pi_b(\mathbf{a}) \geq 1$, it follows from (43) that

$$\sum_{b=1}^B \pi_b(\mathbf{a}) - Br \geq B(1 - r), \quad (44)$$

with equality attained when, for e.g.,

$$a_b = \begin{cases} 0, & b = 1, \dots, \lfloor Br \rfloor \\ \frac{B(1-r)}{B - \lfloor Br \rfloor}, & b = \lfloor Br \rfloor + 1, \dots, B \end{cases} \quad (45)$$

satisfying the constraints in (43). Therefore, $d(r, d_e) \geq mB(1 - r)$ in this case.

- When $B - u - Br > 0$, the right-hand-side of the first constraint in (43) is minimised when, for e.g.,

$$a_b = \begin{cases} 0, & b = 1, \dots, \lfloor Br \rfloor \\ 1 - Br + \lfloor Br \rfloor, & b = \lfloor Br \rfloor + 1 \\ \pi_b(\mathbf{a}), & b = \lfloor Br \rfloor + 2, \dots, B. \end{cases} \quad (46)$$

satisfying the constraints in (42). By letting $a_i^* = a_{i+\lfloor Br \rfloor}$, $i = 1, \dots, B - \lfloor Br \rfloor$, noting that $a_b = a_{b-1} + m \min\{a_{b-u}, d_e\}$, $b \geq u$, we arrive at (16).

Combining the two cases, we arrive at (15). ■

C. Predictive CSIT– Proof of Theorem 2

As in Appendix A-A, for channel with imperfect CSIT $\widehat{\mathbf{H}}^{(\min\{B, b+t\})}$, for large P , the power allocation rule asymptotically satisfies

$$\pi_b \left(\widehat{\omega}^{(b+t)} \right) = 1 + \sum_{b'=1}^{\min\{B, b+t\}} \sum_{i=1}^n (2i-1+m-n) \widehat{\omega}_{b,i}. \quad (47)$$

Therefore, following similar arguments to Appendix A-A, the similar to the arguments in Section A-A, the outage diversity is lower bounded by

$$d(r, d_e) \geq \min_{\mathbf{k} \in \{0, \dots, n\}^B} d_{\mathbf{k}}. \quad (48)$$

In which, $d_{\mathbf{k}}$ is defined as

$$d_{\mathbf{k}} = \inf_{(\overline{\omega}^{(B)}, \widehat{\omega}^{(B)}) \in \mathcal{A}_{\mathbf{k}} \cap \mathcal{O}} \left\{ \sum_{b=1}^B \sum_{i=1}^{n-k_b} (2i-1+m-n) \overline{\omega}_{b,i} + \sum_{i=1}^n (2i-1+m-n) \widehat{\omega}_{b,i} \right\}, \quad (49)$$

where

$$\mathcal{O} = \left\{ \left(\overline{\omega}^{(B)}, \widehat{\omega}^{(B)} \right) : \sum_{b=1}^B \sum_{i=1}^{n-k_b} \left(\pi_b \left(\widehat{\omega}^{(b+t)} \right) - d_e - \overline{\omega}_{b,i} \right)_+ + \sum_{i=n-k_b+1}^n \left(\pi_b \left(\widehat{\omega}^{(b+t)} \right) - \widehat{\omega}_{b,i} \right)_+ < Br \right\} \quad (50)$$

and $\pi_b \left(\widehat{\omega}^{(b+t)} \right)$ is defined in (47). In this case, noting that decreasing $\widehat{\omega}_{b,i}$ decreases both the objective function and the constraint in \mathcal{O} . Therefore, the optimiser satisfies

$$\widehat{\omega}_{b,i} = \begin{cases} d_e, & i \leq n - k_b \\ 0, & \text{otherwise.} \end{cases} \quad (51)$$

and thus $\pi_b \left(\widehat{\omega}^{(b+t)} \right) = 1 + d_e \sum_{b'=1}^{\min\{B, b+t\}} (m - k_{b'}) (n - k_{b'})$. It follows that $d_{\mathbf{k}}$ is obtained from

$$\begin{cases} \text{Infimum} & \sum_{b=1}^B d_e (m - k_b) (n - k_b) + \sum_{b=1}^B \sum_{i=1}^{n-k_b} (2i-1+m-n) \overline{\omega}_{b,i} \\ \text{Subject to} & \sum_{b=1}^B \sum_{i=1}^{n-k_b} \left(\pi_b \left(\widehat{\omega}^{(b+t)} \right) - d_e - \overline{\omega}_{b,i} \right)_+ + \sum_{b=1}^B k_b \pi_b \left(\widehat{\omega}^{(b+t)} \right) < Br \\ & \omega_{b,i} \geq 0, i = 1, \dots, n - k_b \end{cases} \quad (52)$$

as required. ■

APPENDIX B

RDT OF MIMO BLOCK-FADING CHANNELS WITH MISMATCHED CSIT

A. Proof of Theorem 3

Let $h_{b,\tau,v}$ be the fading gain corresponding to the link between transmit antenna τ and receive antenna v in block b . Similarly, let $\hat{h}_{b,\tau,v}$ be the estimate of $h_{b,\tau,v}$ and $\bar{h}_{b,\tau,v}$ be the elements of $\bar{\mathbf{H}}_b$ defined in (3). Further define $\gamma_{b,\tau,v} \triangleq |h_{b,\tau,v}|^2$, $\hat{\gamma}_{b,\tau,v} \triangleq |\hat{h}_{b,\tau,v}|^2$ and $\bar{\gamma}_{b,\tau,v} \triangleq |\bar{h}_{b,\tau,v}|^2$ be the corresponding power fading gains. For an achievability proof, let us consider the power allocation rule

$$\mathbf{P}_b \left(\mathbf{H}^{(b-u)} \right) = P \left(\hat{\mathbf{\Gamma}}^{(b-u)} \right) \mathbf{I}_{N_\tau}, \quad b = 1, \dots, B, \quad (53)$$

where $\hat{\mathbf{\Gamma}} \in \mathbb{R}_+^{B \times N_\tau \times N_v}$ is the power fading gain matrix in block b with entries $\hat{\gamma}_{b,\tau,v}$. Define $\omega_{b,\tau,v} \triangleq \frac{-\log \gamma_{b,\tau,v}}{\log P}$, $\hat{\omega}_{b,\tau,v} = \frac{-\log \hat{\gamma}_{b,\tau,v}}{\log P}$ and $\bar{\omega}_{b,\tau,v} = \frac{-\log \bar{\gamma}_{b,\tau,v}}{\log P}$. It then follows from (1) and (3) that $\bar{\omega}_{b,\tau,v} = \omega_{b,\tau,v} - d_e$. Let $\mathbf{\Omega}_b, \hat{\mathbf{\Omega}}_b, \bar{\mathbf{\Omega}}_b \in \mathbb{R}^{N_\tau \times N_v}$ be the matrices with entries $\omega_{b,\tau,v}, \hat{\omega}_{b,\tau,v}$ and $\bar{\omega}_{b,\tau,v}$ correspondingly. Further define $\pi_b \left(\hat{\mathbf{\Gamma}}^{(b-u)} \right) = \pi_b \left(\hat{\mathbf{\Omega}}^{(b-u)} \right) \triangleq \frac{-\log P \left(\hat{\mathbf{\Gamma}}^{(b-u)} \right)}{\log P}$. Since that the distribution of $\hat{\omega}_{b,\tau,v}$ in the limit of large P is

$$f_{\hat{\omega}_{b,\tau,v}}(\hat{\omega}_{b,\tau,v}) \doteq \begin{cases} P^{-\hat{\omega}_{b,\tau,v}}, & \hat{\omega}_{b,\tau,v} \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (54)$$

it follows from the arguments in Appendix A-A that the power allocation rule with

$$\pi_b \left(\hat{\mathbf{\Omega}}^{(b-u)} \right) = 1 + \sum_{b'=1}^{b-u} \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} \hat{\omega}_{b',\tau,v} \quad (55)$$

is optimal asymptotically.

Following (6) and (8), in the limit of large P , the outage probability is given by

$$\text{P}_{\text{out}}(P, R) \doteq \Pr \left\{ \frac{1}{B} \sum_{b=1}^B I_{\mathcal{X}} \left(\mathbf{H}_b P_b^{\frac{1}{2}} \left(\hat{\mathbf{\Omega}}^{(b-u)} \right) \right) < R \right\}, \quad (56)$$

$$\doteq \Pr \left\{ \sum_{b=1}^B T \left(\hat{\mathbf{\Omega}}^{(b-u)}, \bar{\mathbf{\Omega}}_b \right) > 2^{MN_\tau} B(MN_\tau - R) \right\} \quad (57)$$

where

$$\begin{aligned} T \left(\hat{\mathbf{\Omega}}^{(b-u)}, \bar{\mathbf{\Omega}}_b \right) &\triangleq \sum_{\mathbf{x} \in \mathcal{X}^{N_\tau}} \mathbb{E}_{\mathbf{z}} \left[\log_2 \left(\sum_{\mathbf{x}' \in \mathcal{X}^{N_\tau}} \exp \left(- \left\| \mathbf{H}_b P_b^{\frac{1}{2}} \left(\hat{\mathbf{\Omega}}^{(b-u)} \right) (\mathbf{x} - \mathbf{x}') + \mathbf{z} \right\|^2 + \|\mathbf{z}\|^2 \right) \right) \right] \\ &= \sum_{\mathbf{x} \in \mathcal{X}^{N_\tau}} \mathbb{E}_{\mathbf{z}} \left[\log_2 \left(\sum_{\mathbf{x}' \in \mathcal{X}^{N_\tau}} \exp \left(\sum_{v=1}^{N_v} - \left| \sum_{\tau=1}^{N_\tau} P^{\frac{\pi_b(\hat{\mathbf{\Omega}}^{(b-2)}) - \bar{\omega}_{b,\tau,v} - d_e}{2}} e^{i\theta_{b,\tau,v}} (x_\tau - x'_\tau) \right|^2 + |z_v|^2 \right) \right) \right]. \end{aligned} \quad (58)$$

Here z_v is the v th entry of \mathbf{z} and $i = \sqrt{-1}$ is the imaginary number unit.

For any $\epsilon > 0$, let $\mathcal{S}_b^{(\epsilon)} \triangleq \bigcup_{v=1}^{N_v} \mathcal{S}_{b,v}^{(\epsilon)}$ and $\kappa_b = |\mathcal{S}_b^{(\epsilon)}|$, where

$$\mathcal{S}_{b,v}^{(\epsilon)} = \left\{ \tau \in \{1, \dots, N_\tau\} : \bar{\omega}_{b,\tau,v} + d_e < \pi_b \left(\hat{\Omega}^{(b-u)} \right) - \epsilon \right\} \quad (59)$$

For any $v \in \{1, \dots, N_v\}$, let $\bar{\omega}_{b,v} = \max_{\tau \in \mathcal{S}_{b,v}^{(\epsilon)}} \{\bar{\omega}_{b,\tau,v}\}$. If there exists $\tau \in \mathcal{S}_{b,v}^{(\epsilon)}$ such that $x_\tau \neq x'_\tau$ then

$$\begin{aligned} \lim_{P \rightarrow \infty} \left| \sum_{\tau=1}^{N_\tau} P^{\frac{\pi_b(\hat{\Omega}^{(b-2)}) - \bar{\omega}_{b,\tau,v} - d_e}{2}} e^{i\theta_{b,\tau,v}} (x_\tau - x'_\tau) \right| \\ = \lim_{P \rightarrow \infty} \left| P^{\frac{\pi_b(\hat{\Omega}^{(b-u)}) - \bar{\omega}_{b,v} - d_e}{2}} \sum_{\tau=1}^{N_\tau} P^{\frac{\bar{\omega}_{b,v} - \bar{\omega}_{b,\tau,v}}{2}} e^{i\theta_{b,\tau,v}} (x_\tau - x'_\tau) \right| \\ \geq \lim_{P \rightarrow \infty} \left| P^{\frac{\pi_b(\hat{\Omega}^{(b-u)}) - \bar{\omega}_{b,v} - d_e}{2}} \sum_{\tau \in \mathcal{S}_{b,v}^{(\epsilon)}} e^{i\theta_{b,\tau,v}} (x_\tau - x'_\tau) \right| = \infty \end{aligned} \quad (60)$$

with probability 1 since $\theta_{b,\tau,v}$ are uniformly distributed in $[-\pi, \pi]$. Thus, for asymptotically large P ,

$$T \left(\hat{\Omega}^{(b-u)}, \bar{\Omega}_b \right) \leq \sum_{\mathbf{x} \in \mathcal{X}^{N_\tau}} \log_2 \left(\sum_{\mathbf{x}' \in \mathcal{X}^{N_\tau}} \mathbb{1} \left\{ x_\tau = x'_\tau, \forall \tau \in \mathcal{S}_b^{(\epsilon)} \right\} \right) = 2^{MN_\tau} M (N_\tau - \kappa_b). \quad (61)$$

Therefore, it follows from (57) that

$$P_{\text{out}}(P, R) \leq \Pr \left(\sum_{b=1}^B \kappa_b < \frac{BR}{M} \right) \quad (62)$$

$$\doteq \int_{(\hat{\Omega}^{(B)}, \bar{\Omega}^{(B)}) \in \mathcal{O}} \prod_{b,\tau,v} f_{\bar{\omega}_{b,\tau,v} | \hat{\omega}_{b,\tau,v}}(\bar{\omega}_{b,\tau,v} | \hat{\omega}_{b,\tau,v}) d\bar{\omega}_{b,\tau,v} d\hat{\omega}_{b,\tau,v} \quad (63)$$

where

$$\mathcal{O} = \left\{ \left(\bar{\Omega}^{(B)}, \hat{\Omega}^{(B)} \right) \in \mathbb{R}^{2BN_v N_\tau} : \sum_{b=1}^B \kappa_b < \frac{BR}{M} \right\} \quad (64)$$

$$= \left\{ \left(\bar{\Omega}^{(B)}, \hat{\Omega}^{(B)} \right) : \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \mathbb{1} \left\{ \bar{\omega}_{b,\tau} + d_e < \pi_b \left(\hat{\Omega}^{(b-u)} \right) - \epsilon \right\} < \frac{BR}{M} \right\}. \quad (65)$$

Here $\bar{\omega}_{b,\tau} = \min \{\bar{\omega}_{b,\tau,v}, v = 1, \dots, N_v\}$. Following the arguments in [13], the outage diversity is bounded by

$$d(R, d_e) \geq \inf_{(\bar{\Omega}^{(B)}, \hat{\Omega}^{(B)}) \in \bar{\mathcal{O}}} \left\{ \sum_{(b,\tau,v): -d_e \leq \bar{\omega}_{b,\tau,v} = \hat{\omega}_{b,\tau,v} - d_e < 0} \hat{\omega}_{b,\tau,v} + \sum_{(b,\tau,v): \bar{\omega}_{b,\tau,v} \geq 0, \hat{\omega}_{b,\tau,v} \geq d_e} (\bar{\omega}_{b,\tau,v} + \hat{\omega}_{b,\tau,v}) \right\}, \quad (66)$$

where

$$\overline{\mathcal{O}} \triangleq \left\{ \left(\overline{\boldsymbol{\Omega}}^{(B)}, \widehat{\boldsymbol{\Omega}}^{(B)} \right) \in \mathcal{O} : \{-d_e \leq \overline{\omega}_{b,\tau,v} = \widehat{\omega}_{b,\tau,v} - d_e < 0\} \text{ or } \{\overline{\omega}_{b,\tau,v} \geq 0, \widehat{\omega}_{b,\tau,v} \geq d_e\} \right\} \quad (67)$$

Noting from (67) that for $(\overline{\omega}_{b,\tau,v}, \widehat{\omega}_{b,\tau,v}) \in \overline{\mathcal{O}}$, for (b, τ, v) such that $\widehat{\omega}_{b,\tau,v} \geq d_e$, decreasing $\widehat{\omega}_{b,\tau,v}$ decreases the objective function in (66), while the constraint in (67) is relaxed. Therefore, together with the constraint in (67), the solution of (66) satisfies

$$\widehat{\omega}_{b,\tau,v} = \min \{\overline{\omega}_{b,\tau,v} + d_e, d_e\} \quad (68)$$

Therefore, letting $a_{b,\tau,v} = \overline{\omega}_{b,\tau,v} + d_e$ and $a_{b,\tau} = \min \{a_{b,\tau,v}, v = 1, \dots, N_v\}$, it follows that

$$d(R, d_e) \geq \inf_{\mathbf{a} \in \widehat{\mathcal{O}}} \left\{ \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} a_{b,\tau,v} \right\}, \quad (69)$$

where

$$\widehat{\mathcal{O}} \triangleq \left\{ \mathbf{a} \in \mathbb{R}_+^{B \times N_\tau \times N_v} : \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \mathbf{1}_{\{a_{b,\tau} < \overline{\pi}_b(\mathbf{a}) - \epsilon\}} < \frac{BR}{M} \right\} \quad (70)$$

and $\overline{\pi}_b(\mathbf{a}) = 1 + \sum_{b'=1}^{b-u} \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} \min\{a_{b',\tau,v}, d_e\}$. The infimum in (69) is achievable when, for $v = 1, \dots, N_v$,

$$a_{b,\tau,v} = \begin{cases} 0, & B(b-1) + \tau < \frac{BR}{M} \\ \overline{\pi}_b(\mathbf{a}) - \epsilon, & \text{otherwise.} \end{cases} \quad (71)$$

Therefore, letting $\hat{b} \triangleq \left\lfloor \frac{d_S(R)}{N_\tau} \right\rfloor \triangleq \left\lfloor \frac{1 + \lfloor B(N_\tau - \frac{R}{M}) \rfloor}{N_\tau} \right\rfloor$, the outage diversity is lower bounded by

$$d(R, d_e) \geq N_v N_\tau \sum_{b=1}^{\hat{b}} a'_b + N_v \left(d_S(R) - \hat{b} N_\tau \right) a'_{1+\hat{b}}, \quad (72)$$

where for $v = 1, \dots, N_v$

$$a'_b = \begin{cases} 1 - \epsilon, & b = 1, \dots, u \\ a'_{b-1} + N_\tau N_v \min\{d_e, a'_{b-u}\} - \epsilon, & b = u+1, \dots, \hat{b} + 1. \end{cases} \quad (73)$$

By letting $\epsilon \downarrow 0$, the outage diversity is lower bounded by (19).

On the other hand, using the genie-aided arguments as in [6, 13], the outage diversity is upper bounded by that of a channel consisting of N_τ parallel channels, each is a block-fading channel with N_v receive antenna. Using similar approach as in the previous part of the proof, the outage diversity of the genie-aided channel is also given by (19). This concludes the proof of the Theorem. \blacksquare

B. Proof of Theorem 4

Following the arguments in Appendix B-A, the optimal outage diversity of a MIMO block-fading channel with predictive CSIT $\widehat{\mathbf{H}}^{(b+t)}$ and mismatched CSIT exponent d_e is given by⁵

$$d(R, d_e) = \inf_{\mathbf{a} \in \widehat{\mathcal{O}}} \left\{ \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} a_{b,\tau,v} \right\}, \quad (74)$$

where

$$\widehat{\mathcal{O}} \triangleq \left\{ \mathbf{a} \in \mathbb{R}_+^{BN_\tau} : \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \mathbb{1} \{a_{b,\tau} < \bar{\pi}_b(\mathbf{a})\} < \frac{BR}{M} \right\} \quad (75)$$

$$\bar{\pi}_b(\mathbf{a}) \triangleq 1 + \sum_{b'=1}^{\min\{b+t, B\}} \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} \min \{a_{b',\tau,v}, d_e\}. \quad (76)$$

For any $\mathbf{a} \in \widehat{\mathcal{O}}$, there are $d_S(R)$ coefficients $a_{b,\tau}$'s satisfying $a_{b,\tau} \geq \bar{\pi}_b(\mathbf{a})$. The infimum in (74) is therefore attained when, for $v = 1, \dots, N_v$,

$$a_{b,\tau,v}^* = \begin{cases} \bar{\pi}_b(\mathbf{a}^*), & (b-1)N_\tau + \tau \leq d_S(R) \\ 0, & \text{otherwise.} \end{cases} \quad (77)$$

Then, for b, τ such that $(b-1)N_\tau + \tau \leq d_S(R)$,

$$a_{b,\tau}^* \geq 1 + N_\tau N_v \min \{a_{b,\tau}^*, d_e\} \geq d_e. \quad (78)$$

Thus, when $(b-1)N_\tau + \tau \leq d_S(R)$,

$$a_{b,\tau,v}^* = 1 + \sum_{b'=1}^{b+t} \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} \min \{a_{b',\tau,v}^*, d_e\} \quad (79)$$

$$= \begin{cases} 1 + N_\tau N_v (b+t) d_e, & b+t \leq \left\lfloor \frac{d_S(R)}{N_\tau} \right\rfloor \\ 1 + N_v d_S(R) d_e, & \text{otherwise.} \end{cases} \quad (80)$$

Letting $\hat{b} = \left\lfloor \frac{d_S(R)}{N_\tau} \right\rfloor$. If $\hat{b} \leq t$, the optimal outage diversity is

$$d(R, d_e) = \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} a_{b,\tau,v}^* = N_v d_S(R) (1 + N_v d_S(R) d_e). \quad (81)$$

⁵As in Appendix B-A, the final result is obtained by letting $\epsilon \rightarrow 0$, which does not affect the analysis, hence ϵ is removed for simplicity.

Meanwhile, if $\hat{b} > t$, the optimal outage diversity is

$$\begin{aligned}
 d(R, d_e) &= \sum_{b=1}^B \sum_{\tau=1}^{N_\tau} \sum_{v=1}^{N_v} a_{b,\tau}^* \\
 &= N_v \left(\sum_{b=1}^{\hat{b}-t} \sum_{\tau=1}^{N_\tau} (1 + N_\tau N_v (b+t) d_e) + \left(d_S(R) - N_\tau (\hat{b}-t) \right) (1 + N_v d_S(R) d_e) \right) \\
 &= N_v \left(d_S(R) + N_v d_e \left(\frac{(\hat{b}-t)(\hat{b}+t+1)}{2} N_\tau^2 + d_S(R)(d_S(R) - N_\tau(\hat{b}-t)) \right) \right)
 \end{aligned}$$

as required. ■

REFERENCES

- [1] J. G. Proakis, *Digital Communications*, 5th ed. McGraw Hill, 2008.
- [2] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [3] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
- [4] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Tech.*, vol. 43, no. 2, pp. 359–378, May 1994.
- [5] E. Malkamäki and H. Leib, "Coded diversity on block-fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.
- [6] K. D. Nguyen, "Adaptive transmission for block-fading channels," Ph.D. dissertation, Inst. Telecommun. Research, Univ. South Australia, Dec. 2009. [Online]. Available: <http://www.itr.unisa.edu.au/research/publications/thesis/kdn.pdf>
- [7] R. Knopp and G. Caire, "Power control and beamforming for systems with multiple transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 638–648, Oct. 2002.
- [8] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Trans. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, 2008.
- [9] G. Caire, G. Taricco, and E. Biglieri, "Optimal power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.
- [10] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antenna," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1273–1289, May 2001.
- [11] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, "Outage exponents of block-fading channels with power allocation," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2373–2381, May 2010.
- [12] T. T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, Apr. 2007.
- [13] T. T. Kim, K. D. Nguyen, and A. Guillén i Fàbregas, "Coded modulation with mismatched CSIT over MIMO block-fading channels," *accepted for IEEE Trans. Inf. Theory*, 2010.
- [14] R. Negi and J. M. Cioffi, "Delay-constrained capacity with causal feedback," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2478–2494, Sep. 2002.
- [15] N. Letzepis and A. Guillén i Fàbregas, "Outage probability of the gaussian MIMO free-space optical channel with PPM," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3682–3690, Dec. 2009.

- [16] J. Chen and K.-K. Wong, "Communication with causal CSIT and controlled information outage," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, May 2009.
- [17] K. D. Nguyen, L. K. Rasmussen, A. Guillén i Fàbregas, and N. Letzepis, "MIMO ARQ with multi-bit feedback: Outage analysis," (*submitted to*) *IEEE Trans. Inf. Theory*; available at *arXiv:1006.1162*, May 2009.
- [18] S. V. Hanly and D. N. C. Tse, "Multiaccess fading channels—Part II: Delay-limited capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2816–2831, Nov. 1998.
- [19] L. Zhao, W. Mo, Y. Ma, and Z. Wang, "Diversity and multiplexing tradeoff in general fading channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1549–1557, Apr. 2007.
- [20] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 9, pp. 2632–2639, Sep. 2001.
- [21] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.
- [22] A. Lim and V. K. N. Lau, "On the fundamental tradeoff of spatial diversity and spatial multiplexing of MISO/SIMO links with imperfect CSIT," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 110–117, Jul. 2008.
- [23] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.
- [24] L. Zheng and D. N. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May. 2003.
- [25] H. F. Lu and P. V. Kumar, "A unified construction of space-time codes with optimal rate-diversity tradeoff," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1709–1730, May 2005.
- [26] T. T. Kim and G. Caire, "Diversity gains of power control with noisy CSIT in MIMO channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1618–1626, Apr. 2009.
- [27] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. Springer, 2009.

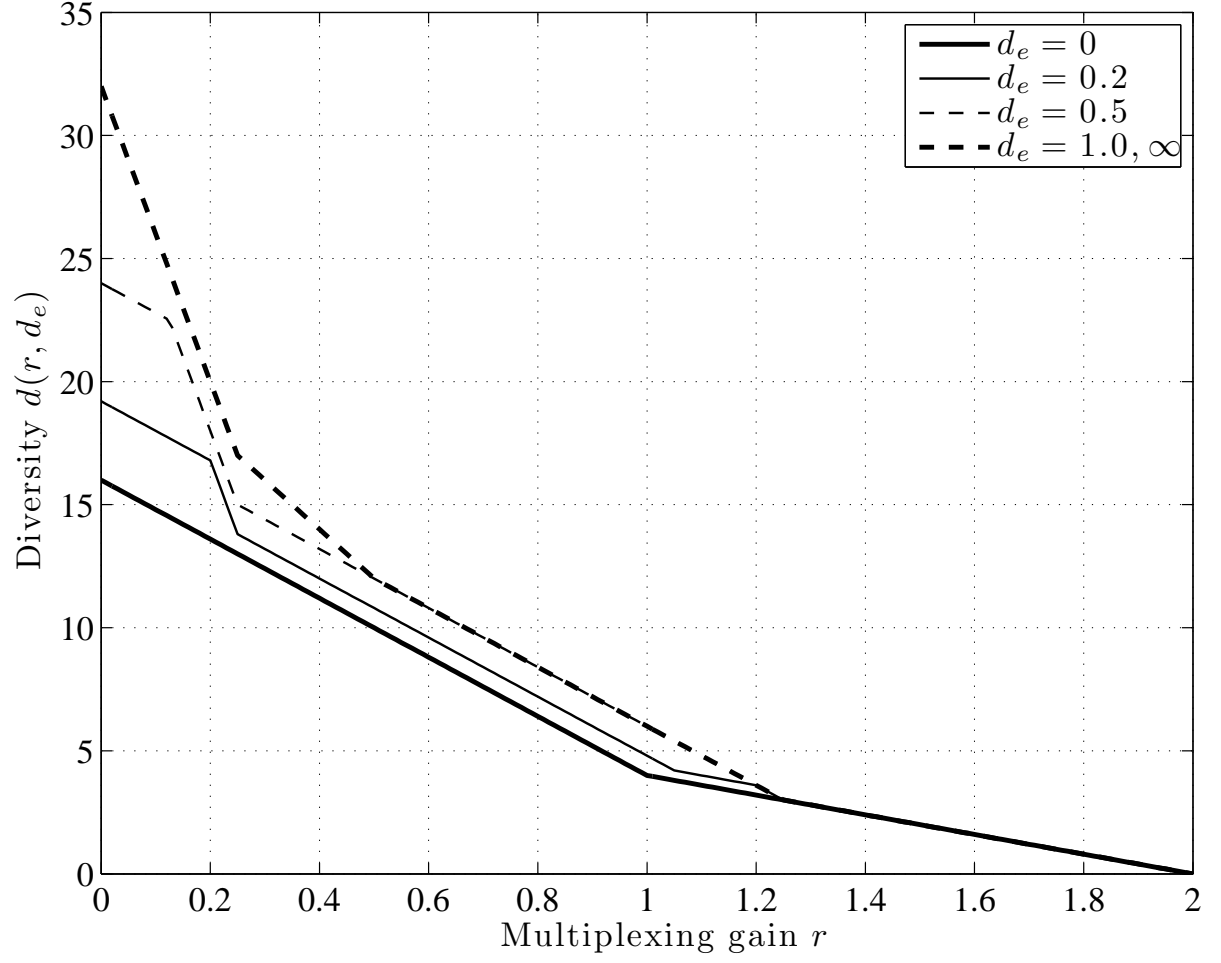


Fig. 1. The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$ causal CSIT $u = 3$.

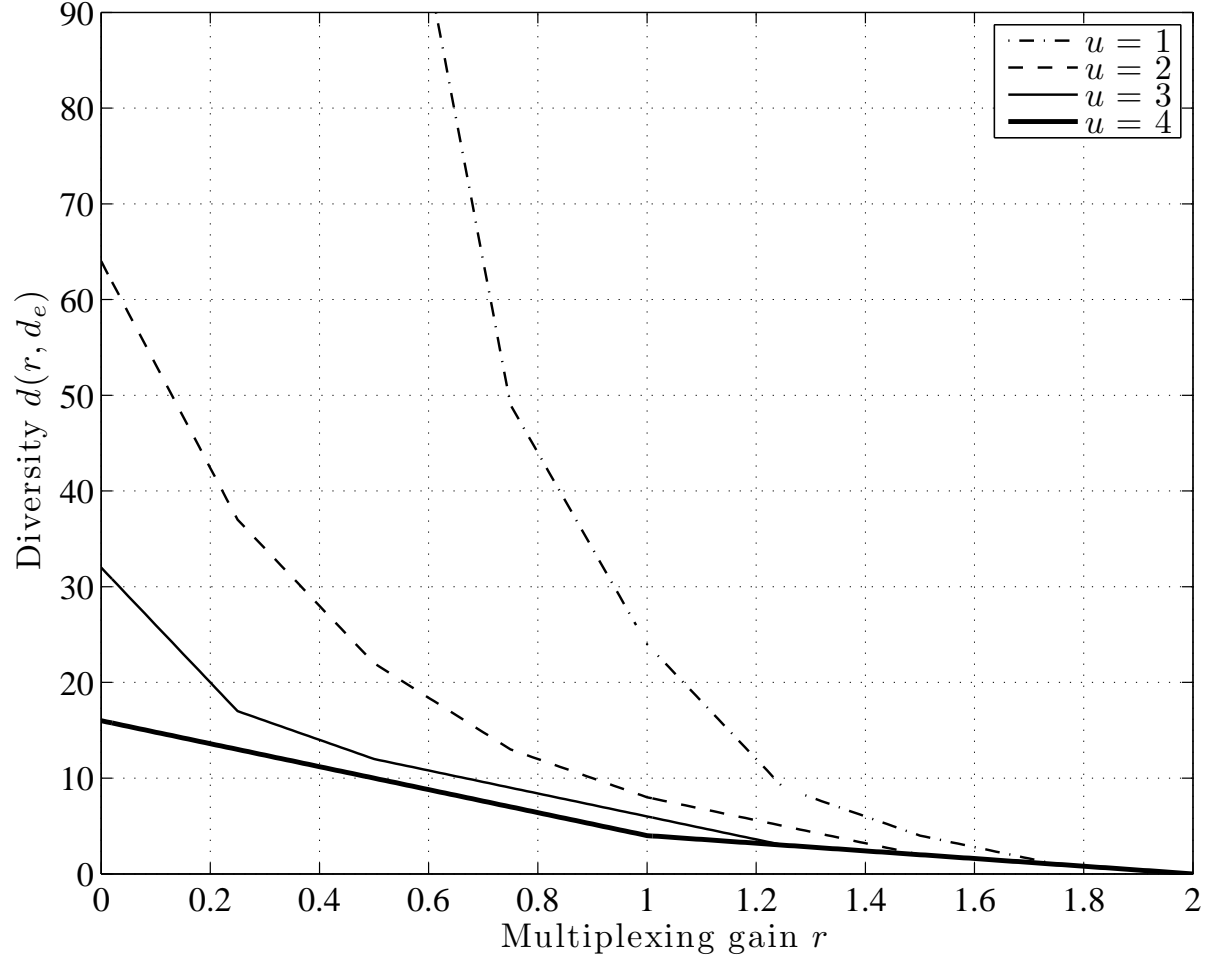


Fig. 2. The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$ and perfect causal CSIT.

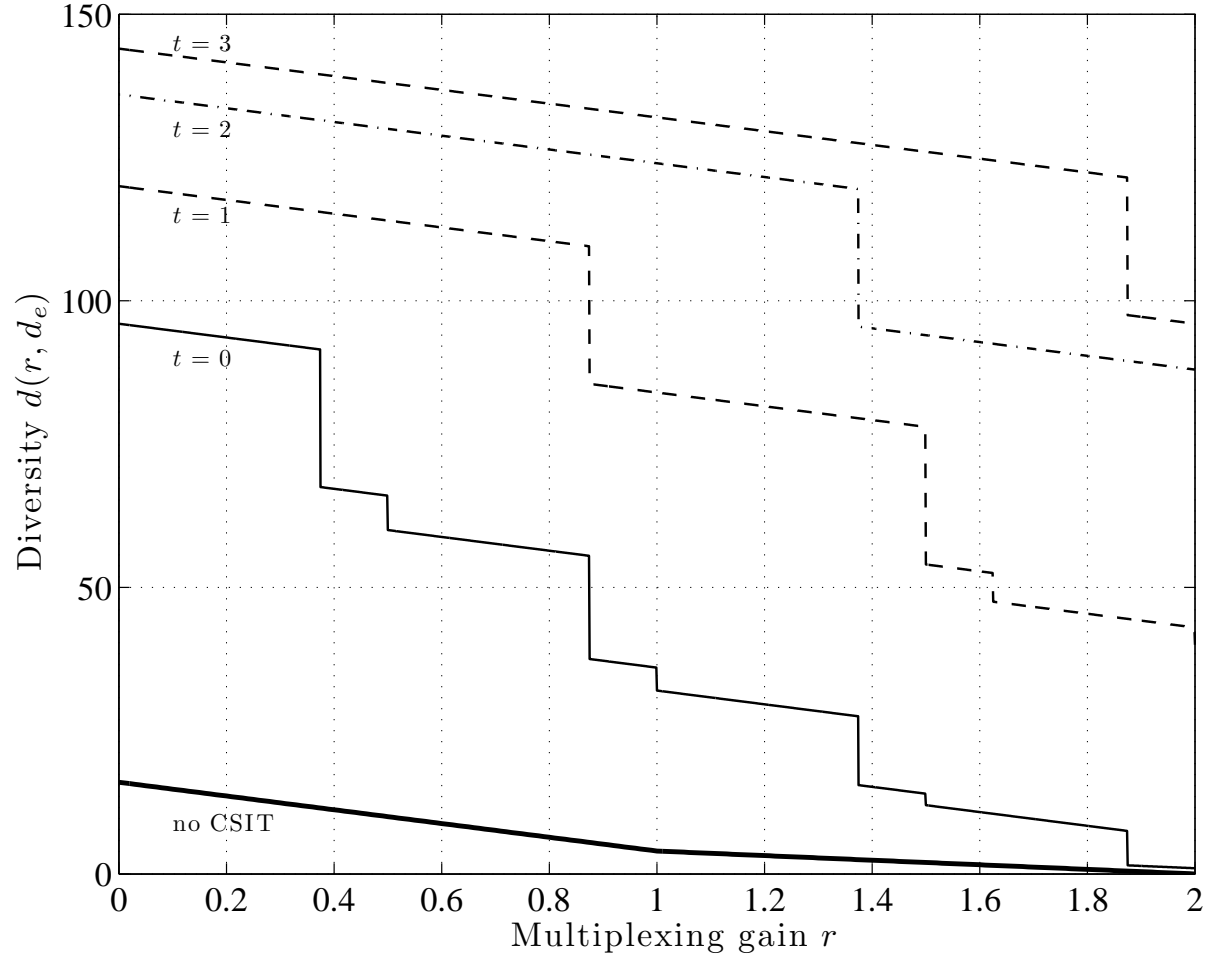


Fig. 3. The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$ predictive mismatched CSIT $d_e = 0.5$

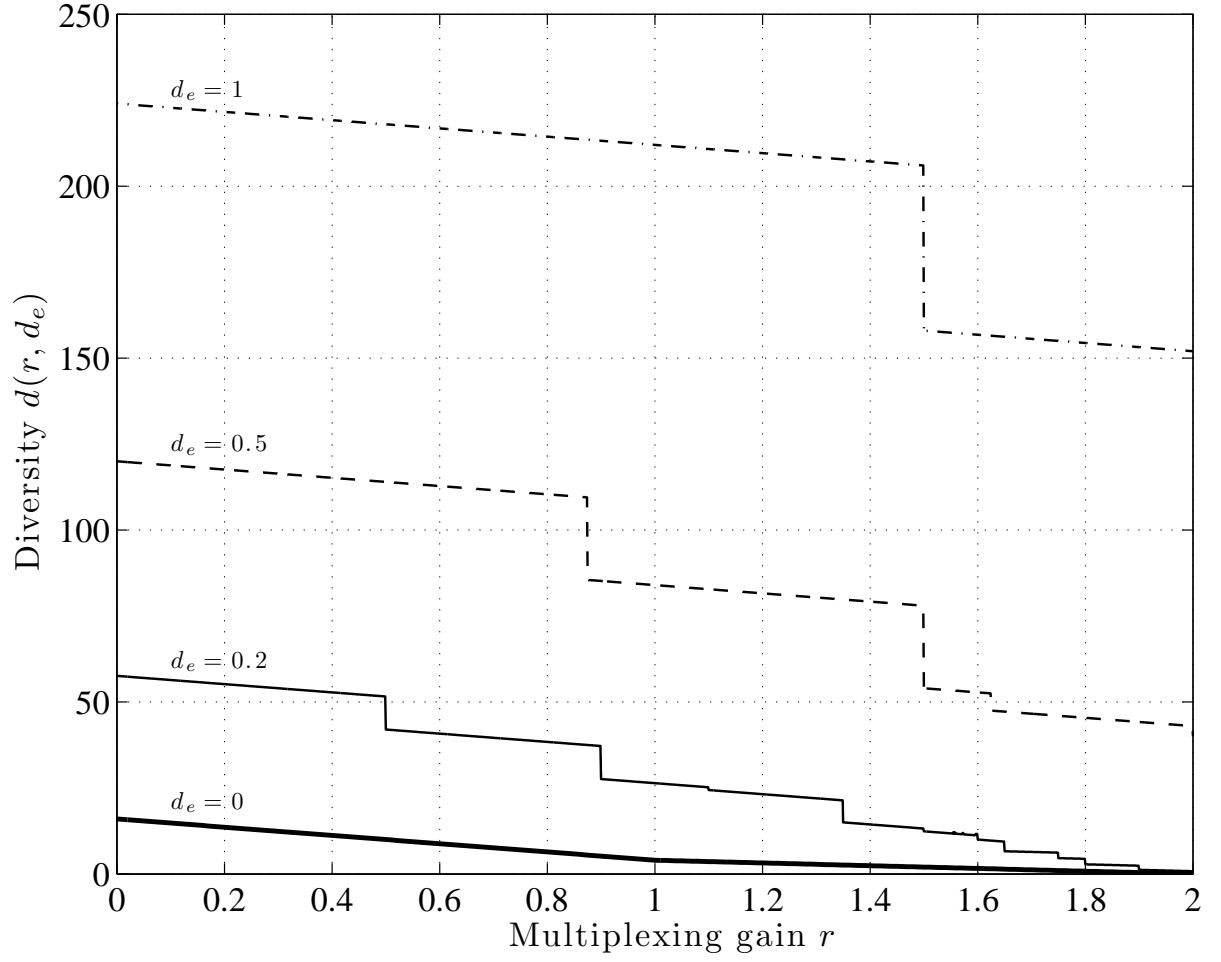


Fig. 4. The achievable DMT of a 2-by-2 MIMO block-fading channel with $B = 4$, predictive mismatched CSIT $t = 1$

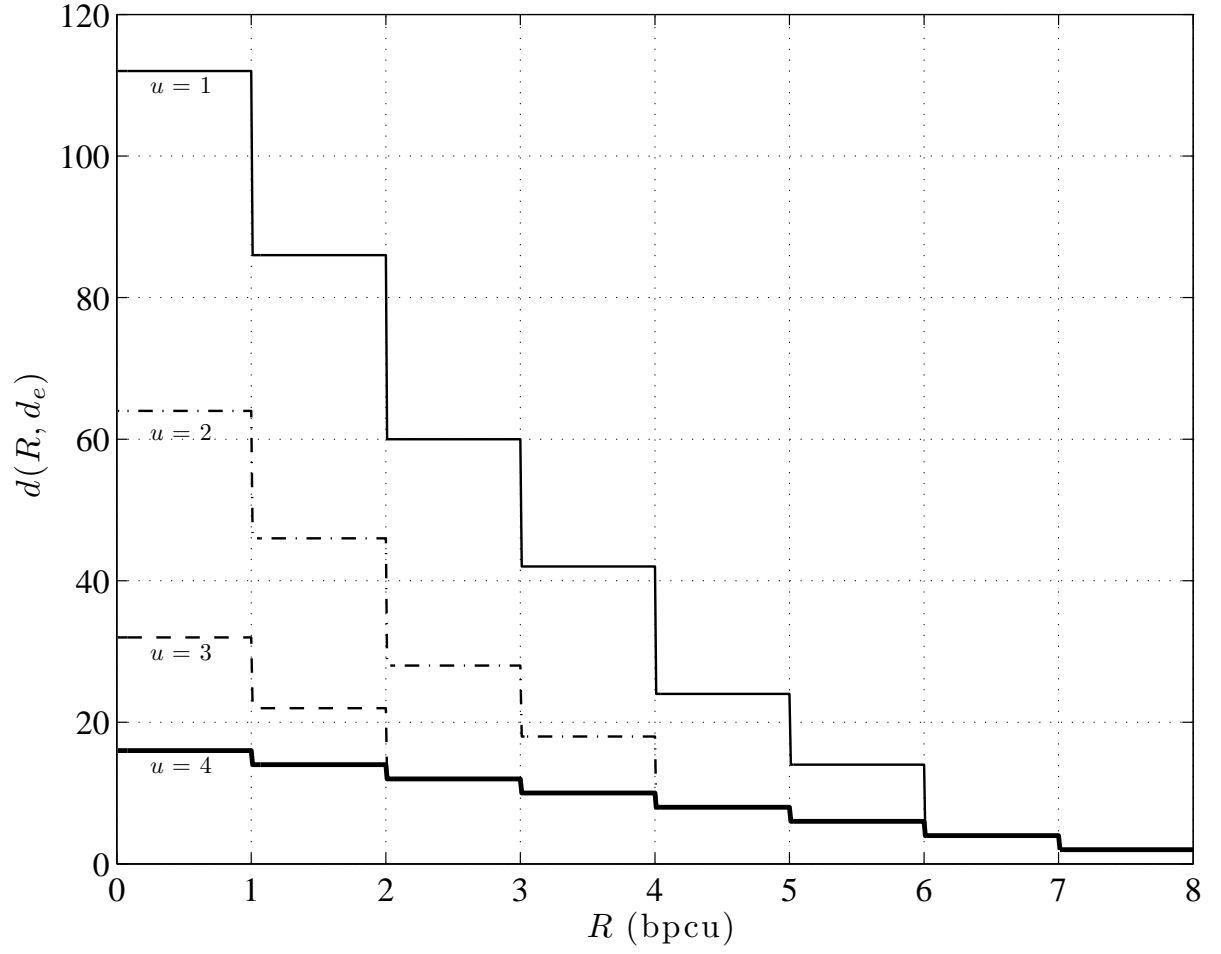


Fig. 5. The optimal rate-diversity tradeoff for a 2-by-2 MIMO block-fading channel with $B = 4$ using 16-QAM input constellation, assuming mismatch causal CSIT with exponent $d_e = 1$.

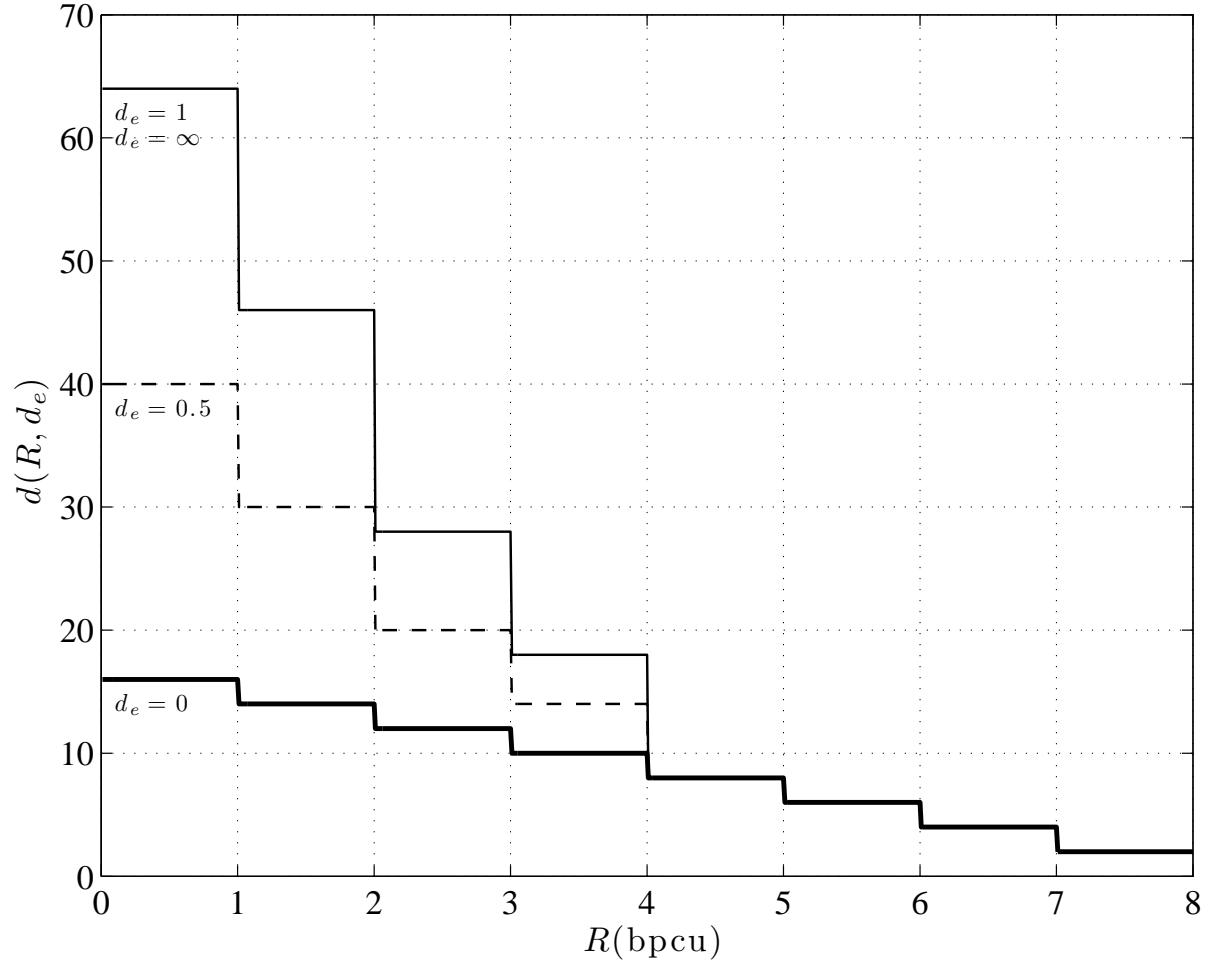


Fig. 6. The optimal rate-diversity tradeoff for a 2-by-2 MIMO block-fading channel with $B = 4$ using 16-QAM input constellation, assuming causal CSIT with delay $u = 2$.

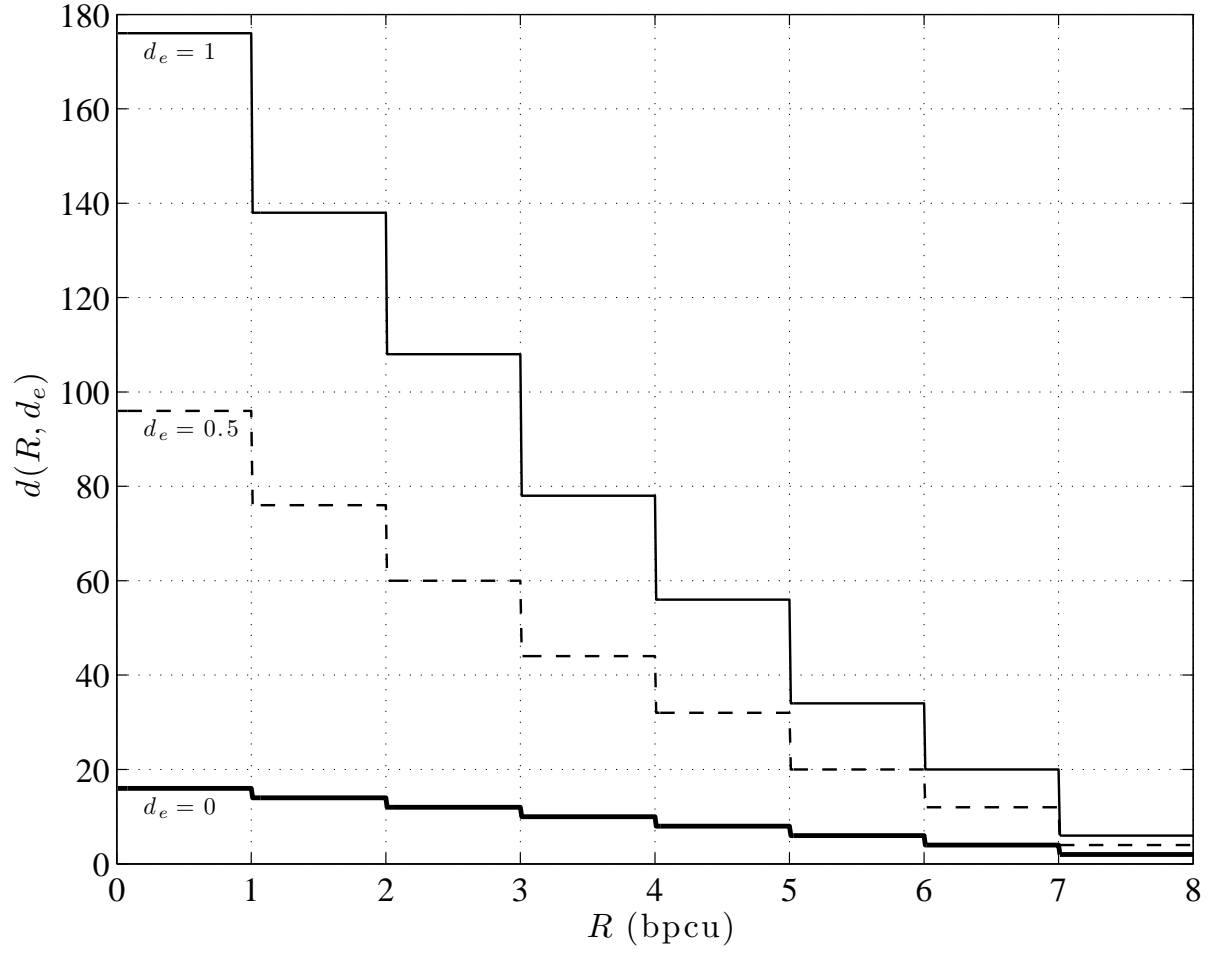


Fig. 7. The optimal rate-diversity tradeoff for a 2-by-2 MIMO block-fading channel with $B = 4$ using 16-QAM input constellation, assuming mismatch predictive CSIT with $t = 0$.

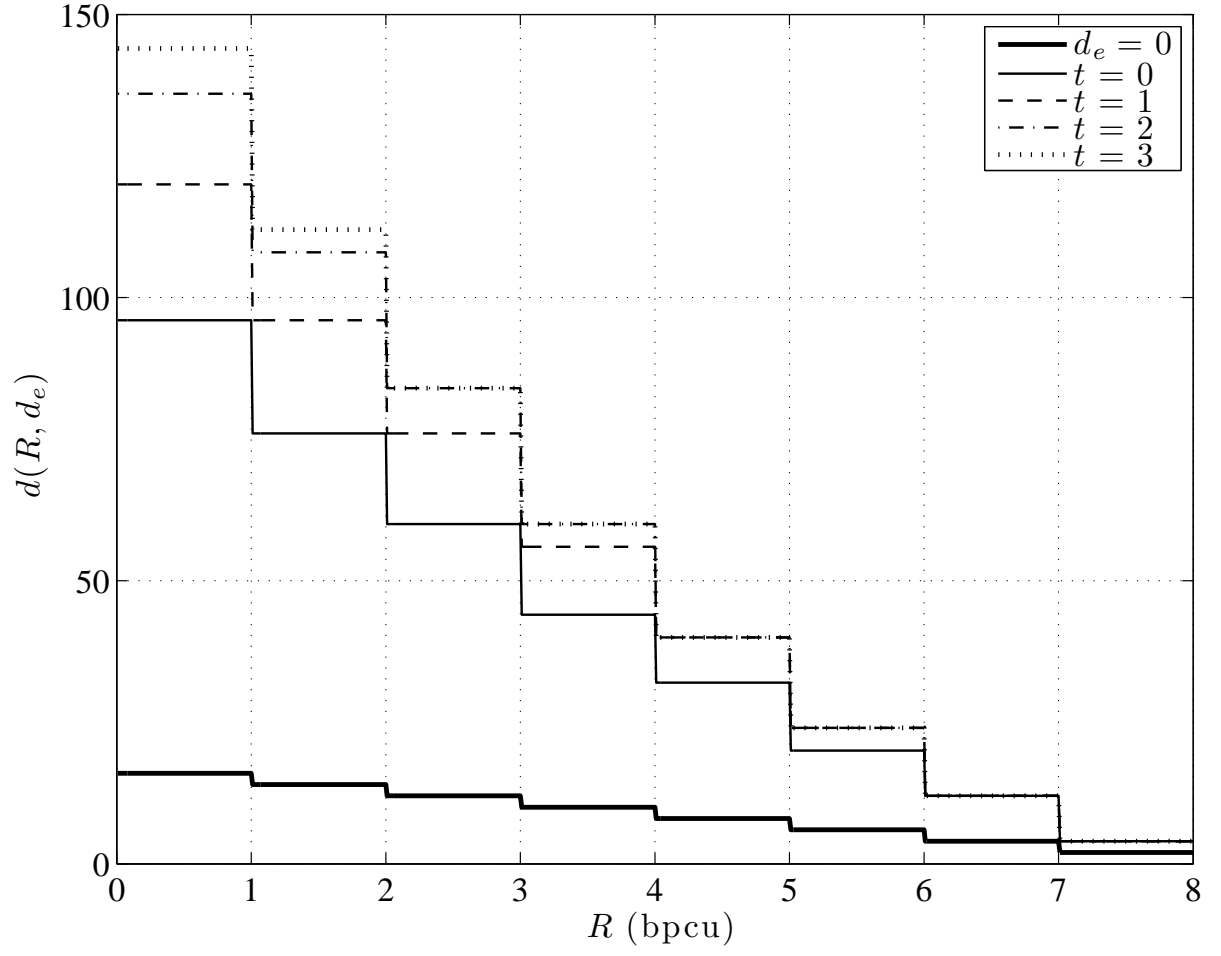


Fig. 8. The optimal rate-diversity tradeoff for a 2-by-2 MIMO block-fading channel with $B = 4$ using 16-QAM input constellation, with predictive CSIT of t block and mismatched exponent $d_e = 0.5$.